END
DATE
FILMED
8-80
DTIC

1.0

28   2.5

3.2   2.2

3.6

4.   2.0

1.1   1.8

1.25   1.4   1.6

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU

LEVEL II

Relating Images, Concepts, and Words*

Working Paper 23

David L. Waltz
Coordinated Science Laboratory and
Department of Electrical Engineering
University of Illinois
Urbana, IL 61801

**DTIC
ELECTE
JUN 2 4 1980
S   D
E**

Abstract

Examination of verbal descriptions of objects suggests that we use hierarchical structures for shape description; the highest levels of the hierarchy provide a general object framework or breakdown into component parts, and a description of each part by analogy to a well-understood set of shapes called prototypes. Lower levels of the hierarchy provide refinement of the analogies and ways in which shapes deviate from the prototypes. The set of prototypes on which the analogies are based contains many common objects, especially natural objects and the parts of the human body, plus certain shapes with special symmetry properties. It is argued that no single 3-D representation scheme is natural for all members of this set of prototypes, and that since unfamiliar objects are described with respect to the basic set of shapes, these objects will have varying shape representation schemes also.

ADA085895

80   5   5   075

Introduction

We perceive objects wherever we look, even when there is very little support for our perceptions. We look at a cluster of stars and see a hunter or a dragon or a dipper. I enter my dark bedroom, and see the heap of blankets on the bed as my wife (example from Minsky 1975). We look at clouds and see dogs or whales or faces. The number of examples could be multiplied manyfold. I suggest in this paper that top-down imposition of objects on weak sensory data is not an isolated, peculiar phenomena, but that most perception proceeds in exactly the same manner, although usually with more reliable sense data, and no conscious awareness of the mapping process.

This paper attempts to provide at least a partial answer to the following questions: (1) How do we represent and describe familiar objects? (2) how do we represent and describe unfamiliar objects? (3) Do we use a uniform representation scheme for all objects? (4) What should the output be for a complete computer vision system? and (5) How can a vision system and a natural language system be integrated and communicate with each other?

The ideas in this paper are a direct result of an investigation into the ways in which objects and parts of objects can be described in natural language. Some examples of the kinds of phrases I encountered are "box canyon", "saw teeth", "table leg", "tail of a kite", "head of lettuce", "clock face", "apple skin", "chain of lakes", and many other similar examples. These

examples could be viewed as "frozen metaphors", but I am struck by the fact that for most of the cases I have looked at, there is no alternate way to refer to the particular object or part of an object.   I therefore have decided to consider the possibility that this kind of apparently metaphorical language might actually be reflecting literal information about the ways that we represent objects, and to see where this assumption leads.

The basic thesis I have developed is this: objects are represented by taking descriptions of well-known prototype objects (or parts of prototype objects) and generating a mapping between these descriptions and unfamiliar objects*. The descriptions of prototype objects are rich, including information about how the object feels, what it looks like from a variety of views, how it can change in shape (if it is non-rigid), and how the object could be composed from components (if it can be). The mapping between familiar and unfamiliar objects allows knowledge of the prototype object to be transferred to the unfamiliar object. Overall, many kinds of mappings are possible, including mappings due to shape similarity ("saw teeth"), similarity of position with respect to the rest of the object ("foot of a tree"), proximity to other objects ("foot of a bed"), and others.

---

*Whenever I refer to an object in this paper, I will mean not only whole objects, such as a human body, but also identifiable parts or components of objects. Thus nose, ear, arm, finger, wheel, doorknob, handle, switch, etc. are meant to be included here when I refer to objects.

I assume throughout that prototype representations are isomorphic in some way to their real world correlates, i.e. non-symbolic (see (Fischler 1978)), and that objects whose shapes are defined in terms of prototypes are thus ultimately grounded in isomorphic representations also. These representations are assumed to aid in the association of image fragments with objects, and to be related to "mental images" (see (Kosslyn 1978)), to visual problem-solving and "mental rotation" of objects, and to the understanding of language which describes physical objects and their relations (Waltz and Boggess 1979). I will first discuss some of my basic notions about the development of perceptual representation schemes, and then go on to describe a number of different kinds of mappings between prototypes and novel objects. I will then give supporting evidence from language, examples of our natural perceptual mapping ability, and some efficiency arguments for the plausibility of these ideas.

Development of perceptual representation schemes

My basic developmental notions are as follows: an infant is born with the ability to form figure/ground relationships, and thus to form concepts (for a discussion of concepts and words, see (Waltz 1978)). However, these concepts have little or no structure or relationship to other concepts. The infant must learn to describe the shape of common objects by painstakingly developing a number of representation schemes for these objects, probably involving constructs such as stick figures, generalized cylinders and cones, surface representations, visual analog and volume

representations, etc. (Marr 1978) and must also learn the typical relationships between objects -- which objects are parts of other objects, which usually occur together, and so on.

One especially important set of objects for an infant to represent consists of the parts of the human body plus the body as a whole. An infant also includes in its object "library" idealizations of real objects such as cubes, spheres, bowls, cylinders, planes, and points, plus representations of many other objects common in the infant's environment, especially natural objects such as trees, birds, fish, and animals (see the discussion in (Bajcsy and Joshi 1978)).

Eventually, once shape representations are developed for a certain number of objects, new objects can be represented much more rapidly and easily by using mappings from, variations on, analogies to, pieces of, and compositions of the shape descriptions already known. The set of objects usually used to describe new objects by analogy thus becomes a "distinguished subset" of "prototype objects" in the terminology of (Winograd 1978).

Processing implications

This paper offers a solution to the object representation problem which is a compromise between extremes: at one extreme is the notion of a single, canonical shape representation scheme suitable for all objects. At another extreme is the search for a set of primitive, complete abstract representations plus methods for finding an appropriate description scheme for representing a

given object. A third extreme (not usually formulated) would be that each object is its own template, with its own unique representation, not necessarily comparable to any other.

In the approach that I am suggesting, representation schemes are initially developed for and attached to particular objects. I also assume that the surface features, silhouette, and possibly other aspects of the objects' appearance are integrated into its representation, so that an association is formed between features of _shape_ and features of _appearance_. This is important, since it allows for a way to select appropriate representations, given sketchy sensory data, and a way to associate tactile features to objects which could never be touched, e.g. jagged mountains or pointed skyscrapers. The method I suggest for selecting representations is roughly:

(1) find an initial 2-d (or 2-1/2-D) segmentation of a scene;

(2) use features with suggestive properties to match prototypes*;

(3) apply prototypes by matching their features with sensory data;

(4) verify the matching on the basis of the properties of adjacent regions (as in (Tenenbaum and Barrow 1975)), or transformations of shape with motion, or functional reasoning, etc.

---

*This assumes that the appearances of the prototype objects from many different perspectives is well-known; however appearances of prototype objects are apparently only well-known for ordinary orientations of the prototype objects - see (Rock 1976).

Later the representations can be applied to new objects (1) by a global mapping (with variations) between the old and new objects, or (2) by taking pieces of the old representation schemes and composing a new representation from the pieces and a framework to "hang" them on. Eventually, certain representation schemes and mapping techniques may be generalized or abstracted, but they would still be ultimately grounded in prototype objects*.

This proposal steers between several difficulties inherent in other approaches: (1) we do not need to assume a canonical shape representation scheme or primitives; (2) we do not need to represent the shapes of all objects in full detail -- we only need to store all the details once, with the prototype objects; (3) the sharing of prototypes and the formation of new representations by variations or compositions of old representations can lead to an overall semantic net-like memory structure with the desirable properties of a natural similarity metric and links for relatedness.

Evidence from language for different kinds of mappings

In this section I present more detailed examples of a number of different ways in which mappings can be formed between prototype objects and sensory data. I will not discuss the issues very completely in this section, but will assume that the examples given

---

*This entire process is reminiscent of Jackendoff's (1975) ideas on the "metaphorical transfer" of schemata for verbs of motion.

in some sense speak for themselves; later I will draw some conclusions about all this.

1. Shape similarity. Many mappings are based on the similarity in shape or topology between one object and another. Some examples of this type of mapping are:

box canyon	armor plate
cupped hands	table rock
brick of cheese/ice	tongue of a shoe
elbow macaroni	pipe elbow
yard arm	fin of an airplane
radiator fins	saw teeth
pipe stem	head of lettuce/cabbage
leaf scale	engine pod
gear teeth	lip of a bowl
light bulb	crotch of a tree
neck of an oar/racquet/bottle/stringed instrument	
brow of a hill	mouth of a river/bottle/cave
dog leg (crooked path)	neck of land
mushroom cloud	funnel cloud
brain coral	star fish
chain of lakes	gold leaf
grease nipple	knuckle coupler
tree/branch/root (data structure, as drawn)	
saddle horn	bell of a trumpet/tuba
pot-bellied stove	stove-pipe hat
beak of a cap	lady fingers
claw of a hammer	rooster's comb
crow's feet	bags under the eyes
shank of a drill/tool	barrel cactus
submarine sandwich	arch of the foot
crotch/limbs/trunk of a tree	

Note that the majority of these examples use the shape of a natural object (part of a person, plant, or animal) to describe the shape of some object, or to denote the part of some object which has the named shape.

2. Position similarity. Often an object part is named by making an analogy between the position of the part relative to the total object and the position of some part of a well-known object

to the whole object. Here are some examples:

            table leg   shoulder of a road
            foot of a mountain   foot of a tree
            head of a river/axe/hammer
            skin of a fruit   tail of a kite/comet/coin
            heart of a city/plant/building/state/country/target
            roots of a hair   skeleton of a building
            rim of a canyon/quarry/crater
            arm of the sea   head land
            cloud ceiling   screw or nail head
            clock face   wasp waist
            flank of a hill   crest of a hill

    3.   Proximity to other objects. Sometimes objects or parts of

objects are named by reference to the parts of other objects which

are usually in close proximity to them. Examples:

            foot of a bed   head of a bed
            toe of a boot   heel of a boot
            neck of a sweater   finger of a glove
            waist of a dress or trousers
            mouthpiece of a wind instrument
            elbow pad   eye glasses
            handle (hand-le)   pedal (ped-al; ped=foot in Latin)
            headstone   earphones
            throat microphone   hip pocket
            <any of a large number of objects> + cover
            corner store   beach house
            door bell   foot locker
            chair back   back of a coat/shirt/jacket
            hand rail

    When objects are used in phrases to modify other objects which

are containers, a special kind of proximity (enclosure and often

support) is conveyed:

            silverware drawer   briefcase (brief case)
            flower pot   coffee cup
            soup bowl   ice tray
            jewelry box   perfume bottle
            cereal box   candle holder
            medicine cabinet   ice house
            car barn   fish bowl

Certain objects have a proximity relationship which is interpreted as meaning <u>covering</u>. Phrases of the form <object>+cloth/cover usually have this meaning. Examples:

tablecloth	loin cloth
wall paper	face mask
car cover	bedclothes
book cover	food wrap
skull cap	bottle cap

Still other objects suggest proximity relationships which include <u>support</u>. Examples:

coat/hat rack	spice rack
TV stand/table	bookshelf
coffee table	coat hook
coat hanger	picture hanger
telephone receiver	plant stand
dog bed	automobile jack/lift
roof pillar	antenna mast
light pole	telephone pole
foot bridge	railroad bridge

4. Objects with marked orientation. To quote (Rock 1976), "...the perception of form embodies an automatic assignment of a top, a bottom and sides." Many objects have by convention a inherent front, back, top, bottom, and sides. These objects in a sense have had a cube with a marked front mapped onto them. Furthermore all objects, even those without inherent marked orientations, can be assigned a front, back, top, and so on by reference to a viewer's position (e.g. the <u>front of</u> the tree") or by reference to some other object (e.g. "the side of the mountain <u>facing</u> Pompeii"). The various parts of objects which have marked orientation (whether inherent or assigned automatically) can be referred to by terms like top, bottom, side, front, back, head,

tail, and base. Some examples:

```
car (headlight, taillight, side molding, rear door, front
 wheels,convertible top, undercoat, back bumper, etc.)
house (front/side/back doors/windows)
animals (front legs, back legs, sides, head, tail)
people (frontal nudity, back, sides, tops of head/shoulders)
desk (top drawer, front, sides, back, bottom drawer)
```

Virtually any object which has some marked (non-symmetric) axis, and which usually appears in some preferred orientation, can be assigned this type of mapping; consider airplanes, TV sets, books, buildings, phones, boats, bottles (with labels), stoves, clocks, vacuum cleaners, blossoms, chairs, etc.

As discussed in (Clark 1973), objects which can be seen by a speaker or listener can be assigned a marked direction (front, back, top, etc.) even if the objects do not have any inherent marked orientation.   Thus we can say that a ball is in back of a tree, meaning that from where the speaker is standing, the tree has a front which faces the speaker, and a back which faces away form the speaker. Clark calls this a "canonical encounter" and suggests that objects are treated as though they were people being met face-to-face.

Sometimes we can take the point of view of the listener when we are speaking, as when one might say to a seeker in a game of hide and seek "I'm hiding in back of the tree." Sometimes the canonical encounter coordinate system gives a different assignment of front, back, etc. from the inherent coordinate system of an object with marked direction, and meaning ambiguity results; thus "The ball is behind the car" could mean either that the ball is in

back of the rear bumper, or that (if the car were being viewed from the side) that the ball is on the side of the car opposite the speaker (see Figure 1). Sometimes we can distinguish the inherent orientation of the object by using "the", as in "The fly is on top of the bottle" vs. "The fly is on The top of the bottle." In the former case, the fly could simply be on the highest part of a bottle, which might be the side if the bottle were lying down, whereas in the latter case, the fly could only be on the part of the bottle near its mouth (see Figure 2).

Larger scale analogies

Certain analogies involve the simultaneous mapping of a number of parts between objects. The most extensive example I have encountered is that of an airplane where the parts are described with respect to a bird. Thus an airplane has a tail, wings, belly, skin, and skeleton/frame; it also has other parts described in a kind of "mixed metaphor," namely its nose, radio antenna, tail fin, and engine pod. Ships also seem to be described by a similar large scale analogy with an animal: a ship has a nose, belly, ribs, tail, and skin.

Cartoons and drawings in children's books also provide examples of large scale mappings. Cartoon animal characters are often created by mapping animal heads, feet, and tails onto the heads, feet and rumps of a prototype human body. Consider Bugs Bunny, Donald Duck, Mickey Mouse, the Big Bad Wolf, etc., etc. All walk upright, have roughly human proportions, have human facial
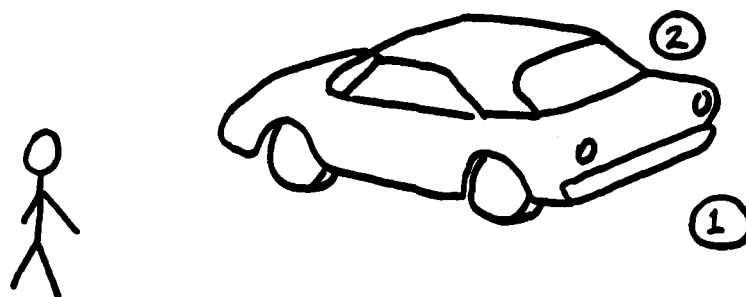
Figure 1 Two interpretations of
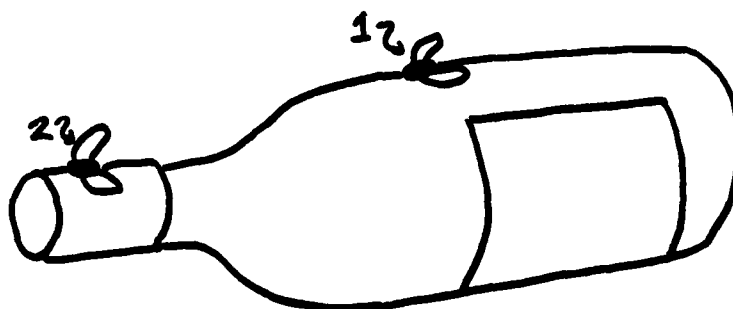"The ball is behind the car."

Figure 2 "The fly is on top of the bottle" (1) vs.
"The fly is on the top of the bottle" (2).

characteristics added, e.g. eyebrows, and so on (a detailed analysis of Donald Duck is given below). As shown in figure 3, human characteristics can also be mapped onto less obvious candidates with relatively little detailed similarity, e.g. human features onto airplanes, trains, cars, boats, houses, mountains, trees, and so on. In each case just described there is one prototype object which provides the framework onto which other prototype objects are mapped. Thus in the cases of Donald Duck et al, a human body provides the framework, and animal body parts are mapped onto and attached to the human framework. In the case of figure 3, human faces have been mapped onto the frameworks of an airplane and a mountain peak. These kinds of mapping may have interesting relationships to the notion of animism (Piaget 1967), i.e. the universal childhood propensity to view inanimate objects as animate agents with goals and intentions. Some examples include the very frequent conviction on the part of children that the moon follows them as they walk, and the universal addition of eyes, nose and mouth to drawings of the sun.

Surrealist art also provides some interesting connections with these ideas. For example, consider figures 4, 5, and 6; figure 4 shows a reverse mermaid (fish from the waist up and woman from the waist down - see (Minsky 1975) for a discussion of this in terms of frames), and in figure 5, the features of a human face have been replaced by vegetables. Figure 6 is a "portrait" of Mae West, which on closer inspection is actually a room with furniture. This painting is particularly striking in that it shows how readily a specific face can be mapped onto a set of objects having the right

Figure 3 Human faces mapped onto an
airplane and a mountain from the story
"Pedro" in Walt Disney's Treasury, Golden
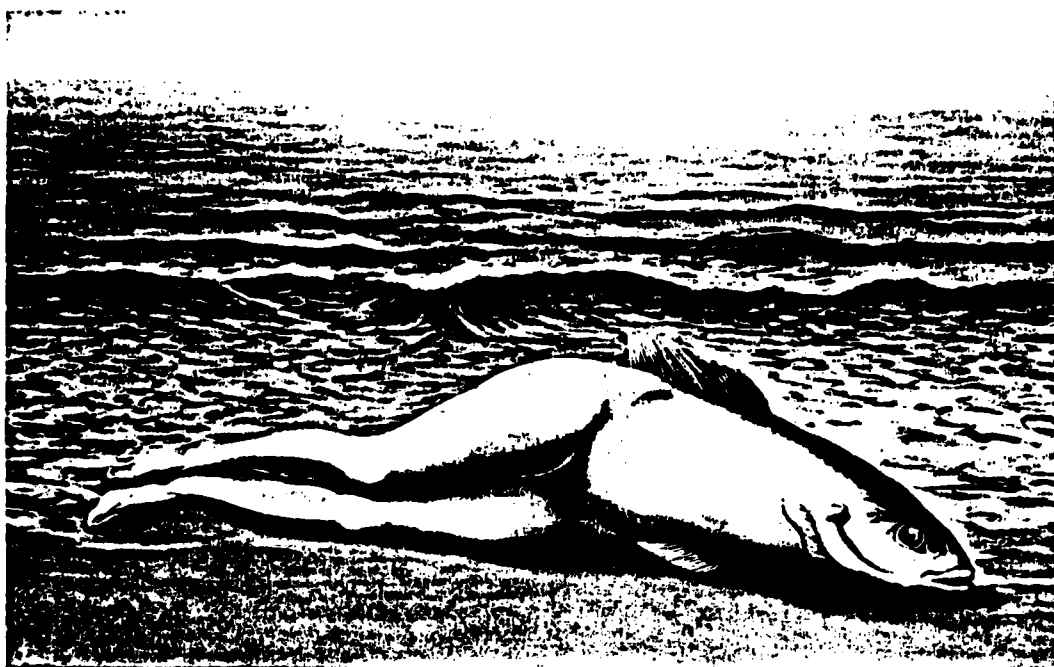Press, New York, 1953.

Figure 4 "Collective invention" (1935) by Rene Magritte. From Suzi Gablik, Magritte, New York Graphic Society Ltd., Greewich, Connecticut, 1970.

Figure 5 "The Market Gardener" by Guiseppe Arcimboldo, 16th century. From Sarane Alexandrian, Surrealist Art, Praeger, New York, 1970.

Figure 6 "Mae West" (1934-6) by Salvador Dali.
From Sarane Alexandrian, Surrealist Art, Praeger,
New York, 1970.

2-D arrangement; presumably, it is even easier to map a <u>general</u> prototype onto an image. There are numerous other examples in surrealist art of this kind of playing with the objects that are hung on a familiar framework: fur covered cups, a nude woman whose body is partially flesh and partially wood grain, clouds in the shapes of a tuba, chair and torso, and so on.

The fact that English uses a very similar set of words to describe the parts of people, mammals, reptiles, insects, and fish, and that these terms date from long before the theory of evolution, suggests that we are inclined to make analogies between objects and their parts, and to thereby economize on words, even when the feature-by-feature shape similarity is slight (as between the noses of a variety of animals). To a certain degree, such mappings may also reflect matching of parts by functional rather than shape similarity or similarity of position with respect to the whole organism.

Mapping

I have not yet defined precisely what I mean by the term <u>mapping</u>; the following description is sketchy, but should at least give an idea of what I have in mind. Mappings are of two main types: <u>structural</u> and <u>topological</u>.

By structural mapping, I mean roughly that both objects being related by the map can be described by abstracted structures, e.g. stick figures or graphs, and that components of the two abstracted structures can be associated to form the map. Examples of

structural mapping are the part-by-part association of a person's body with a chimpanzee's body, or the association of the markings on a pansy blossom with the eyes, nose, and mouth of a person's face.

By topological mapping, I mean something more like deformation or coordinated system transformation, which allows points on the surface of one object to be associated with points on the surface of the other object. Examples of topological mappings are the duck's head to a sphere mapping mentioned above, the "Cartesian transformations" of (D'Arcy Thompson 1969) -- see figure 7 -- or a mapping of an object such as a mountain or a piece of a saw blade onto a prototype "tooth", or the mapping of a cube onto an arbitrary object (as in the examples of assigning front, sides, top, etc. to objects).

I assume that structural mapping should precede topological mapping, and may be used as a kind of filter for testing whether a more detailed topological mapping is feasible. Topological mapping is the only kind possible for relatively structureless objects like spheres, and may involve intermediate level representations such as "shape envelopes" of objects, i.e. the surface shape of objects with the detail suppressed (see (Waltz 1978b) for some ideas on finding shape envelopes of 2-D objects). Much difficult mathematical work remains to be done here!
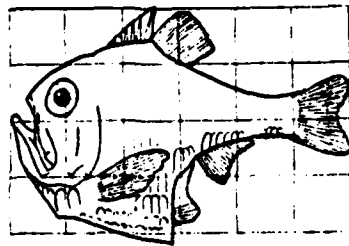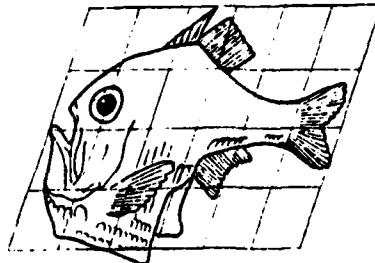
Assessment

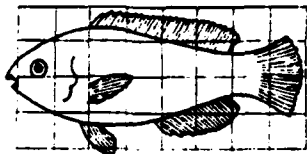Fig. 146. *Argyropelecus olfersi.*

Fig. 147. *Sternoptyx diaphana.*
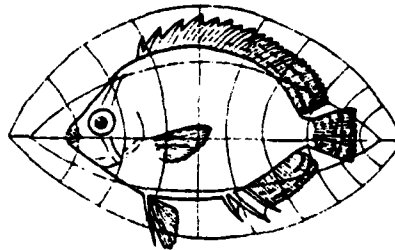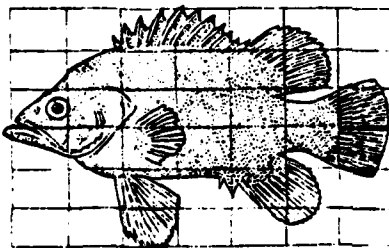
Fig. 148. *Scarus* sp.

Fig. 149. *Pomacanthus.*
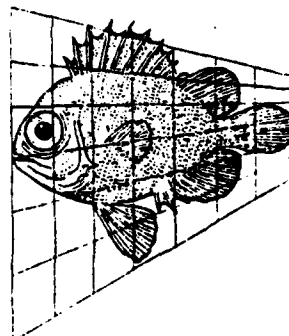
Fig. 150. *Polyprion.*

Fig. 151. *Pseudopriacanthus altus.*

Figure 7 Examples of mappings by "Cartesian
transformations" from D'Arcy Thompson, On Growth
and Form, Cambridge University Press, 1961.

The words and phrases of English provide support for the idea that objects are represented as combinations and variations on prototypes. However, the evidence is like archaeological evidence, in that the word descriptions are not invented by each language user, but are given to each of us as part of our cultural heritage. The descriptions are often reminders of the kinds of objects (mostly natural) which were available to describe artificial objects when they were first introduced. For a child learning to speak today, there is no reason to suppose that a bird is any more familiar than an airplane -- language may serve to encourage a child to make an analogy between the two (a la (Whorf 1956)), but both objects are probably represented in some manner independently before this happens. What we can say is that when a totally unfamiliar object is encountered (e.g. an airplane to people in 1903) the tendency is to see the unfamiliar object as analogous to well-known objects, and to describe the parts of the unfamiliar object using the vocabulary of familiar objects. The types of analogy made are also noteworthy; analogies are most naturally made for objects with similar frameworks and similar shapes. We do not as readily make analogies between objects based on functional similarity (train and airplane are both modes of transportation, but share relatively little as objects), or similarity of material, or frequent cooccurrence, or other possible similarities. Perhaps this seems self-evident, but let me drive home the point that object shape seems to be the most important factor in naming or describing objects.

There is also evidence that people are good at and naturally do generate mappings from familiar to novel objects. For example, consider the process of learning to identify all the things we call faces as instances of the concept face. Children must learn to deal with this very broad sensory category by developing a representation scheme which judges all sensory items in the category face to be similar. I suggest that the natural representation for similarity is what I have called a prototype, and that it is a 3-D visual shape analog representaion.

I feel that this outline is plausible by arguments of efficiency alone: different objects (e.g. ball, human body, table, spoon, cup, box) are most naturally described by quite different representations schemes*.
Once an infant has developed representation schemes for describing a sufficiently large set of objects, new objects seldom require that new representation schemes be developed; old schemes can with relatively less effort be applied to the new objects. Eventually the set of objects for which structures have already been constructed becomes large enough so that new objects do not require that the representation schemes be used at all; instead, part or all of the representation structure itself from some old object will fit a new object (or part of the new object) sufficiently well so that only minor modifications of the old structure plus a mapping between objects is necessary to describe the new object.

---

*I wish to include in "representation scheme" both a target structure (e.g. a graph or generalized cone) plus procedures for generating the structure.

A further efficiency argument can be made for the use of analogy for object description: in addition to describing the shape of objects (probably integrating tactile and visual information) an infant also learns to recognize the objects from many diffeent perspectives, and thus at least implicitly, an infant understands the transformations of appearance of a given shape under rotation. This knowledge of transformations of appearance can be transferred to new objects by analogy, and can also be used in constructing the analogy to begin with. For example, once an infant can easily recognize a coin in any orientaton, he or she can guess that an apparent ellipse might really correspond to a circular coin-like shape.

In a similar manner, dynamic properties of objects such as their behavior when flexed, pressed, bent, dropped, scratched, cut, and so on can also be transferred from prototype objects. Similarities is dynamic object behavior may lead to categories such as rigid/nonrigid, solid/plastic/liquid or animate/inanimate (see (Pylyshyn 1977)). These categories are orthogonal to <u>static</u> shape, but are clearly important for understanding shape transformations.

Problems remaining

Clearly a great deal of work is needed before the ideas in this paper will be a practical part of a vision system. Special problems include picking a set of prototype objects, developing schemes for mapping and composing representations, developing methods for indexing the prototype from image features, developing

appropriate similarity metrics and measuring procedures, and so on. Moreover, suitable low-level vision systems must be developed to provide the kind of image representation which can function with this higher-level vision system.

The particular scheme argued for here has been developed with the conviction that it is dangerous to study vision (or language) in isolation; the function of vision is to organize the sensory data from an eye into a conceptual structure which one can reason about, describe in language, or operate on (e.g. through a manipulator)*. The main effort here is to suggest a plausible higher-level vision system to begin with. In my estimation, inadequate thought has been given to the problem of describing a total vision system; few people have even worried about what the output of a total vision system should be, and few have written about how the piece of a system they are programming (e.g. for segmentation) might fit into a complete system.

It also seems clear to me that we must develop better methods

---

*The study of language in isolation has led to notions that are very dubious, e.g. that the solving of anaphoric reference problems should be done by heuristic search through the series of parse trees generated by the sentences in a dialogue or text. I would argue instead that language is much more closely related to picture-building (Fillmore 1977, Talmy 1978) and that the solving of anaphoric reference has more similarity to scene understanding than to heuristic search.

for dealing with the problems of matching, analogy formation, mapping and structure transfer, for many reasons other than the ones I have discussed in the body of this paper. We "see" complete objects even when the objects are partially occluded or oriented away from us. We can judge how objects will fit together (e.g. puzzles, model car parts, etc.), where objects will break if stressed, how to cut away material to make a given shape from a block, and whether two objects in different orientations are similar. All these operations seem to require matching, mapping, and verification processes (although much more would be needed as well).

Moreover, in the long run, I believe that abstract thought is possible only by metaphorical transfer of schemata from the sensory/motor world to a series of other worlds which may eventually have very little contact indeed with the physical world. Such transfers depend on having a rich, well-developed set of representations for the physical world from which to map to other worlds, and on having good matching, analogy-making, and structure mapping facilities available.

REFERENCES

[1]   R. Bajcsy and A.K. Joshi, "The Problem of Naming Shapes: Vision-Language Interface," in D.L. Waltz (Ed.), TINLAP-2: Theoretical Issues in Natural Language Processing-2, Association for Computing Machinery and The Association for Computational Linguistics, 1978, pp. 157-161.

[2]   H.H. Clark, "Space, Time, Semantics, and the Child," in T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language, New York: Academic Press, 1973.

[3]   C.J. Fillmore, "The Case for Case Reopened," in J.P. Kimball (Ed.), Syntax and Semantics, New York: Academic Press, 1977.

[4]   M.A. Fischler, "On the Representation of Natural Scenes," in A.R. Hanson and E.M. Riseman (Eds.), Computer Vision Systems, New York: Academic Press, 1978, pp. 47-52.

[5]   R. Jackendoff, "A System of Semantic Primitives," in R. Schank and B. Nash-Webber (Eds.), Theoretical Issues in Natural Language Processing, ACL, Arlington, Virginia, 1975.

[6]   S.M. Kosslyn and S.P. Shwartz, "Visual Images as Spatial Representations in Active Memory," in A.R. Hanson and E.M. Riseman (Eds.) Computer Vision Systems, New York: Academic Press, 1978, pp. 223-242.

[7]   D. Marr, "Representing Visual Information," in A.R. Hanson and E.M. Riseman (Eds.) Computer Vision Systems, New York: Academic Press, 1978, pp. 61-80.

[8]   G.A. Miller, and P. Johnson-Laird, Language and Perception, Cambridge, MA: Harvard University Press, 1976.

[9]   M. Minsky, "A Framework for Representing Knowledge," in P.H. Winston (Ed.) The Psychology of Computer Vision, New York: McGraw-Hill Book Co., 1975, pp. 211-277.

[10] J. Piaget, Six Psychological Studies, New York: Vintage Books, 1967.

[11] Z.W. Pylyshyn, "Children's Internal Descriptions," in Language Learning and Thought, New York: Academic Press, 1977, pp. 169-176.

[12] I. Rock, "The Perception of Disoriented Figures," in Scientific American, Jan. 1974, Vol. 230, No. 1, pp. 78-86.

[13] L. Talmy, "The Relation of Grammar to Cognition - a Synopsis," in D.L. Waltz (Ed.) TINLAP-2: Theoretical Issues in Natural Language Processing-2, Association for Computing Machinery and The Association for Computational Linguistics, 1978, pp. 14-24.

[14] J.M. Tenenbaum and H.G. Barrow, "Experiments in Interpretation-Guided Segmentation," <u>Artificial Intelligence</u>, Vol. 8, No. 3, 1977, pp. 241-274.

[15] D.W. Thompson, <u>On Growth and Form</u>, London: Cambridge University Press, 1969.

[16] D.L. Waltz, "On the Interdependence of Language and Perception," in D.L. Waltz (Ed.), <u>TINLAP-2: Theoretical Issues in Natural Language Processing-2</u>, Association for Computing Machinery and The Association for Computational Linguistics, 1978, pp. 149-156.

[17] D.L. Waltz, "A Parallel Model for Low-Level Vision," in A.R. Hanson and E.M. Riseman (Eds.) <u>Computer Vision Systems</u>, New York: Academic Press, 1978, pp. 175-186.

[18] D.L. Waltz and L. Boggess, Working Paper 20: "Visual Analog Representations for Natural Language Understanding."

[19] B.L. Whorf, <u>Language,Thought and Reality</u>, Cambridge: The M.I.T. Press, 1956.

[20] T. Winogard, "On Primitives, Prototypes, and Other Semantic Anomalies," in D.L. Waltz (Ed.) <u>TINLAP-2: Theoretical Issues in Natural Language Processing-2</u>, Association for Computing Machinery and The Association for Computational Linguistics, 1978, pp. 25-32.

DATE
ILMED

8